

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 9, 2003, 06:27:29 ; Search time 17 Seconds
(without alignments)

919.799 Million cell updates/sec

Title: US-10-091-628-2
 Perfect score: 1979
 Sequence: 1 MRANSSSSACPNASSEEL.....PGMDCHRALEPVGHITSCE 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

112892 segs, 41476328 residues
 Total number of hits satisfying chosen parameters: 112892

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	886	44.8	347	1	NTCI1_RABIT	Q28717 c ileal soc
2	884	44.7	348	1	NTCI1_CRIGR	Q60414 c ileal soc
3	871	44.0	348	1	NTCI1_RAT	Q62633 r ileal soc
4	860.5	43.5	348	1	NTCI1_HUMAN	Q12908 h ileal soc
5	559.5	28.3	362	1	NTCP_RAT	P26435 ratus norv
6	553	27.9	349	1	NTCP_HUMAN	Q14973 homo sapien
7	553	27.9	362	1	NTCP_MOUSE	Q08705 mus musculu
8	333.5	16.9	477	1	P3_HUMAN	P09131 homo sapien
9	180.5	9.1	182	1	P3_MOUSE	P112129 mus musculu
10	125	6.3	409	1	YCCA_BACSU	Q08791 bacillus su
11	117	5.9	721	1	YCIY_ECOLI	Q09396 escherichia
12	116	5.9	286	1	YCXE_BACME	P40419 bacillus me
13	114	5.8	372	1	OPRD_HUMAN	P41143 homo sapien
14	112.5	5.7	368	1	CTB_TOXCO	Q20672 toxoplasma
15	109.5	5.5	383	1	Y944_SYNY3	P74311 synchocystis
16	107.5	5.4	576	1	NYSM_ANOQU	P33510 anopheles c
17	106.5	5.4	443	1	COX8_HUMAN	Q12887 homo sapien
18	104.5	5.3	397	1	YGED_ECOLI	P39156 escherichia
19	104	5.3	551	1	LDDP_ECO57	Q8x4d9 escherichia
20	104	5.3	551	1	LDDP_ECOLI	P33231 escherichia
21	101.5	5.1	532	1	YCI8_HAEIN	P57251 haemophilus
22	100.5	5.1	332	1	YFEH_ECOLI	P39836 escherichia
23	99.5	5.0	379	1	CVB_BOLM	P21722 bolomyes ameo
24	99.5	5.0	387	1	GAL5_HUMAN	Q43603 homo sapien
25	99.5	5.0	436	1	SECY_METJA	P060175 methanococc
26	99.5	5.0	541	1	NUSM_ARSPF	Q37710 artemia salin
27	99	5.0	312	1	CCSA_ODOSI	P49523 odontella s
28	99	5.0	333	1	YK26_PPSAE	P39879 pseudomonas
29	98.5	5.0	391	1	SOTB_HELIP	Q9zK31 helicobacter
30	98	5.0	391	1	YIDY_ECOLI	P1462 escherichia
31	98	5.0	835	1	CD97_HUMAN	P48960 homo sapien
32	97.5	4.9	460	1	YICJ_ECOLI	P31435 escherichia
33	97	4.9	369	1	MRAY_AANSP	Q8yp33 arabidra sp

34	96.5	4.9	404	1	SOTB_PASWU	Q9cm47	
35	96	4.9	444	1	YQJA_BACSU	O9cm62	pasturella
36	96	4.9	465	1	DCDR_XENIA	O13682	baetiellus su
37	96	4.9	521	1	FZD7_MOUSE	P42291	xenopus lae
38	95.5	4.8	521	1	FZD7_SALTI	O61090	mus musculu
39	95.5	4.8	551	1	LBDP_SALTI	O82263	salmonella
40	95	4.8	551	1	LBDP_SALTI	O82163	salmonella
41	95	4.8	347	1	NU2M_CANPA	O92655	canis famila
42	95	4.8	380	1	CYB_ONCMY	P48173	oncorhynch
43	95	4.8	380	1	CYB_SALSA	O35925	salmo salar
44	95	4.8	380	1	CYB_SALTR	P29671	salmo trutta
45	95	4.8	551	1	COXI_APIII	P20374	apis mellif
46	95	4.8	556	1	HYRF_ECOLI	P77437	escherichia

ALIGNMENTS

ID	NTCI_RABIT	STANDARD;	PRT;	347 AA.
AC	Q28727;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Ileal sodium/bile acid cotransporter (Ileal Na(+)/bile acid cotransporter) (Na+ dependent ileal bile acid transporter) (Ileal sodium-dependent bile acid transporter) (ISBT) (Sodium/caurocholate cotransporting polypeptide, ileal).			
DE	SLC10A2 OR NTC2.			
GN	Oryctolagus cuniculus (Rabbit).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=New Zealand white; TISSUE=ileum;			
RA	Stengelin S., Abel S., Becker W., Meier M., Rosenberger J., Wess G., Kramer W.;			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: PLAYS A CRITICAL ROLE IN THE SODIUM-DEPENDENT REABSORPTION OF BILE ACIDS FROM THE LUMEN OF THE SMALL INTESTINE.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO THE SODIUM:BILE ACID SYMPORTER FAMILY (SBP).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).			
CC	-----			
DR	EMBL; Z54357; CAA91184.1; .			
DR	EMBL; AJ002005; CAA05135.1; .			
DR	InterPro; IPR004710; Bass.			
DR	InterPro; IPR002657; BileAc/Na_smprtc.			
DR	Pfam; PF01758; SBP; 1.			
DR	TIGRFam5; TIGR00841; Bass; 1.			
RW	Transmembrane; Transport; Symport; Sodium transporter; Glycoprotein.			
FT	TRANSMEM 30 50 POTENTIAL.			
FT	TRANSMEM 59 79 POTENTIAL.			
FT	TRANSMEM 83 103 POTENTIAL.			
FT	TRANSMEM 128 148 POTENTIAL.			
FT	TRANSMEM 159 179 POTENTIAL.			
FT	TRANSMEM 197 217 POTENTIAL.			
FT	TRANSMEM 226 246 POTENTIAL.			
FT	TRANSMEM 290 310 POTENTIAL.			
FT	CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT	CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).			
SO	SEQUENCE 347 AA; 37729 MW; 1315B6BADDE66C CRC64;			

Query Match 44.8%; Score 886; DB 1; Length 347;

FT TAUCHOLATE TRANSPORT.
 FT /FTid=VAR_004615.
 FT P -> S (IN CD; ABOLISHES TAUCHOLATE
 FT TRANSPORT).
 FT /FTid=VAR_004616.
 FT /FTid=VAR_004616.
 SQ SEQUENCE 348 AA; 34897 MM; 159900AAA91CCD06 CRC64;
 Query Match 43.5%; Score 860.5; DB 1; Length 348;
 Best Local Similarity 45.6%; Pred. No. 2,7e-49;
 Matches 160; Conservative 58; Mismatches 104; Indels 19; Gaps 4;

QY 5 CSSSSACPRASSEELPVGLEVHGNLEIVTVSTVMGMLMSLGCSEVIRLMSHIR 64
 DB 14 CCGAGCVCVPBSNNNI-----LSVSLVTLTLTLALVMSWGCNVEIKFELGHIRK 64
 QY 65 PMGIAGVLLCOFGLMFTAYLLAISFSLKPVQAVLIMCCPGGTISNIFFWVDGMD 124
 DB 65 PMGICVGLCOFGLMFTAYLLAISFSLKPVQAVLIMCCPGGTISNIFFWVDGMD 124
 QY 125 LLSMTTCSTVALGMPPLCIYLYTWSMSLQONLTPYQNIITLVCLIPVAFGVYVNY 184
 DB 125 LLSVMTTCSTVALGMPPLCIYLYTWSMSLQONLTPYQNIITLVCLIPVAFGVYVNY 184
 QY 185 RMPKSKITLIGAVVGVLLVAVAGVVLAKGSNNSDITLTISFIRPLIGHVTFLL 244
 DB 185 RMPKSKITLIGAVVGVLLVAVAGVVLAKGSNNSDITLTISFIRPLIGHVTFLL 244
 QY 245 ALFTHSWQRCRTISLETGAQNIQMCITMLQSLFTEHVLVOMLSPFLAYGLFOLLIDGFLI 304
 DB 245 ALFTHSWQRCRTISLETGAQNIQMCITMLQSLFTEHVLVOMLSPFLAYGLFOLLIDGFLI 304
 QY 245 ARIAGLPWYRCRIVAVETGQNTQSLTIVQSLFTEBELNVVTFPLIISIFOLARFAIF 304
 DB 245 ARIAGLPWYRCRIVAVETGQNTQSLTIVQSLFTEBELNVVTFPLIISIFOLARFAIF 304
 QY 305 VAAVQYKRLKRNKHKKSGCTEVCHTRKSTSSRETNAFLVNEGAIIP 355
 DB 305 VAAVQYKRLKRNKHKKSGCTEVCHTRKSTSSRETNAFLVNEGAIIP 355
 QY 305 LGFYVAYKK-----CHGKNKAEIPE-----SKENCTEPESSFYKAN--GGFOP 345
 DB 305 LGFYVAYKK-----CHGKNKAEIPE-----SKENCTEPESSFYKAN--GGFOP 345

RESULT 5
 NTCIP_HUMAN STANDARD; PRT; 362 AA.
 ID NTCIP_HUMAN STANDARD; PRT; 362 AA.
 AC P26435;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Sodium/bile acid cotransporter (Na(+)/bile acid cotransporter)
 DE (Na(+)/taurocholate cotransport protein) (Sodium/taurocholate
 DE cotransporting polypeptide).
 GN SLC10A1 OR NTCIP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92073340; PubMed=1961729;
 RX Hagenbuch B., Stieger B., Foguet M., Luebert H., Meier P.J.;
 RX Functional expression, cloning and characterization of the hepatocyte
 RX Na(+)/bile acid cotransporter system.
 CC Proc. Natl. Acad. Sci. U.S.A. 88:10629-10633 (1991).
 CC FUNCTION: THE HEPATIC SODIUM/BILE ACID UPTAKE SYSTEM EXHIBITS
 CC BROAD SUBSTRATE SPECIFICITY & TRANSPORTS VARIOUS NONBILE ACID
 CC ORGANIC COMPOUNDS AS WELL. IT IS STRICTLY DEPENDENT ON THE
 CC EXTRACELLULAR PRESENCE OF SODIUM.
 CC SUBCELLULAR LOCATION: Integral membrane protein.
 CC TISSUE SPECIFICITY: LIVER AND KIDNEY.
 CC SIMILARITY: BELONGS TO THE SODIUM/BILE ACID SYMPORTER FAMILY.
 CC (SBR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M77479; AAA42112.1; -
 CC DR PIR: A41601; A41601.
 CC DR InterPro: IPR004710; Baas.
 CC DR InterPro: IPR002657; BileAc/Na_smptr.
 CC DR Pfam: PF01758; SBR; 1.
 CC DR TIGR/Pfam: TIGR00841; Baas; 1.
 CC KW TRANSMEMBRANE; Transport; Symport; Sodium transport; Glycoprotein.
 FT TRANSMEM 24 45
 FT TRANSMEM 60 80
 FT TRANSMEM 82 98
 FT TRANSMEM 158 178
 FT TRANSMEM 190 211
 FT TRANSMEM 228 244
 FT TRANSMEM 285 306
 FT TRANSMEM 5 5
 FT CARBOHYD 103 111
 FT CARBOHYD 113 113
 FT CARBOHYD 117 117
 FT CARBOHYD 271 271
 SQ SEQUENCE 362 AA; 36295 MM; FOAB76076A57550 CRC64;
 Query Match 28.3%; Score 559.5; DB 1; Length 362;
 Best Local Similarity 37.2%; Pred. No. 9.7e-30;
 Matches 133; Conservative 69; Mismatches 135; Indels 21; Gaps 9;

QY 10 ACPANSEEDLVGLEVHGNLEIVTVSTVMGMLMSLGCSEVIRLMSHIRPWCIA 69
 DB 7 SAPNFNS---LPPGCG-HRADKALSTIIVMLLMISLGTMTFSKIKALMKPKVGI 62
 QY 70 VGLLCOFGLMFTAYLLAISFSLKPVQAVLIMCCPGGTISNIFFWVDGMDLISJM 129
 DB 63 VALVAGFGLMFTAYLLAISFSLKPVQAVLIMCCPGGTISNIFFWVDGMDLISJM 122
 QY 130 TTCSSTFSAAGMPLLLVYSKGIYDGLDKR--VYKGMISLVLPCTIGIVLKSRR 180
 DB 123 TTCSSTFSAAGMPLLLVYSKGIYDGLDKR--VYKGMISLVLPCTIGIVLKSRR 180
 QY 187 PKOSKILIKIGAVVGVLLVAVAGVVLAKGSNNSDITLTISFIRPLIGHVTFLL 244
 DB 181 PHVYVYLKGMPLIIFLVSANVTALSVNGSINFWMTPHLANSIMPSGFLMGITL 240
 QY 245 -ALFTHSWQRCRTISLETGAQNIQMCITMLQSLFTEHVLVOMLSPFLAYGLFOLLIDGFLI 302
 DB 241 SALF--QINPSCRTISMETGFQNIQSLTILNVTPEPEVIGPLFFPLIYMIFOLARGL 298
 QY 303 LIVAAYQYKRLKRNKHKKSGCTEVCHTRKSTSSRETNAFLVNEGAIIPGPPGP 360
 DB 299 LIIIFRCYEKI-----KPPKDYKITYKAATEDATPALEKCTHNGINIPLOPPG 350

RESULT 6
 NTCIP_HUMAN STANDARD; PRT; 349 AA.
 ID NTCIP_HUMAN STANDARD; PRT; 349 AA.
 AC Q14973;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sodium/bile acid cotransporter (Na(+)/bile acid cotransporter)
 DE (Na(+)/taurocholate cotransport protein) (Sodium/taurocholate
 DE cotransporting polypeptide).
 GN SLC10A1 OR NTCIP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94179485; PubMed=8132774;
 RX Hagenbuch B., Meier P.J.;

RESULT 7	
NTCP_MOUSE	
ID_NTCP_MOUSE	STANDARD;
	PRT;
	362 AA

```

AC 008705; 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sodium/bile acid cotransporter (Na(+)/bile acid cotransporter)
DE (Na(+)/taurocholate transporter protein) (Sodium/taurocholate
DE cotransporting polypeptide).
GN SLCT0A1 OR NTCF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Sasaki T.;
RN Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Hagenbuch B.;
RN Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: THE HEPATIC SODIUM/BILE ACID UPTAKE SYSTEM EXHIBITS
CC BROAD SUBSTRATE SPECIFICITY & TRANSPORTS VARIOUS NONBILE ACID
CC ORGANIC COMPOUNDS AS WELL. IT IS STRICTLY DEPENDENT ON THE
CC EXTRACELLULAR PRESENCE OF SODIUM.
CC -! SUBCELLULAR LOCATION: Integral membrane protein.
CC -! SIMILARITY: BELONGS TO THE SODIUM:BILE ACID SYMPORTER FAMILY
CC (SBF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AB003303; BAA19846.1; -.
DR EMBL; U95131; AAB81023.1; -.
DR MGD; MG1:97379; SLC10a1.
DR InterPro; IPR0004710; Bass.
DR InterPro; IPR002657; BileAc/Na_symprt.
DR Pfam; PF01758; SBF; 1.
DR Trasmembrane; TIGR00841; Bass; 1.
FT TRANSMEM 24 45 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 82 98 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 190 211 POTENTIAL.
FT TRANSMEM 228 244 POTENTIAL.
FT TRANSMEM 285 306 POTENTIAL.
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 362 AA; 39413 MW; 7A70493EB1804280F CMC64;

Query Match 27.9%; Score 553; DB 1; Length 362;
Best Local Similarity 35.1%; Pred. No. 2.6e-29;
Matches 129; Conservativity 72; Mismatches 142; Indels 24; Gaps 8;

QY 10 ACPANSSSEELPYGLEVGHNLVFTVSTVMGMLMPSLGCSSYIRKLWGHIRPWCIA 69
DB 7 SAPNFNS---LPPEFG-HRAIDTALSYLVVMLILMISLCCETFSKIKAHFWKPKYVI 62
QY 70 VGLLQCFGLMFTAYLLAISFLKPVQAAVILMGCCPGGTISNIFPMVGDMDLSISM 129
DB 63 IAIYAQYIMPLSAFLILGKVFHLSIEKALILICGCSFGNLSNFLAMGDMNLSIVM 122
QY 130 TTCTGVAALGMPPLCIYLT--MSWSLQONULTIPYQNIIGTIVCLTIPVAFGYVYVNR 186

```

```

Db      123  TTCSTFTALGMMPLLLIYTSKGIYDGLDKD--VPEYKGMISLWVLLIPCAIGIFLKSXR 180
QY      187  PKQSKIILKIGAVVCGVLLLVAVAGVVLAKSNNSDIT--LITISFIFPLIGHATGYLL 244
Db      181  PNYVYVYLKAGMIITFSLLSVAATVLSVINGNSIMFYMTPLHATSSLMFPGLMGYIL 240
QY      245  ALFTHQSNORCRSTISLETGAONICMTIMLQLSFTAEHLVQMLSPFLAYGIFQLIDGFLI 304
Db      241  SALPFLNLSCRTIMETGSPQVQVQGLSTLANTLPPEVIGLFFPFLPLVIMFQLAEGFLF 300
QY      305  VAAVYTTYRRLLKNKFKGKKNKSGCTEVYCHTRKSTSSRETNAPLEVNEEGAIITPPGPPMDCH 364
Db      301  IIFPCY--LKIKPKQDQ---TKITVYKAAATADATPAALEKGTINGNPNPTPG----- 349
QY      365  RALEPVG 371
Db      350  --LSPNG 354

RESULT 8
HUMAN
P3_HUMAN          STANDARD;          PRT;          477 AA.
PO5331
01-MAR-1989 (Rel. 10, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P3 protein.
GN      P3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxId=9606;
RX      MEDLINE=89041548; PubMed=3186440;
RX      MEDLINE=96311563; PubMed=8733135;
RA      Chen E.Y., Zollo M., Mazarella R.A., Ciccodicola A., Chen C.-N.,
RA      Zhuo L., Heiner C., Burrough F.W., Ripetto M., Schlessinger D.,
RA      D'Urso M.;
RT      "Long-range sequence analysis in Xq28: thirteen known and six
RT      candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and
RT      G6PD loci.";
RL      Nucleic Acids Res. 16:9527-9556(1988).
[2]
SEQUENCE FROM N.A.
RP      MEDLINE=96311563; PubMed=8733135;
RA      Chen E.Y., Zollo M., Mazarella R.A., Ciccodicola A., Chen C.-N.,
RA      Zhuo L., Heiner C., Burrough F.W., Ripetto M., Schlessinger D.,
RA      D'Urso M.;
RT      "Long-range sequence analysis in Xq28: thirteen known and six
RT      candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and
RT      G6PD loci.";
RL      Hum. Mol. Genet. 5:659-668(1996).
-1- FUNCTION: THE UBIQUITOUS EXPRESSION AND THE CONSERVATION OF THE
SEQUENCE IN DISTANT ANIMAL SPECIES SUGGEST THAT THE GENE CODES FOR
A PROTEIN WITH HOUSEKEEPING FUNCTIONS.
-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
-1- SIMILARITY: TO P3 PROTEIN OF ANIMALS AND YEASTS.
-1- SIMILARITY: BELONGS TO THE SODIUM:BILE ACID SYMPORTER FAMILY
(SBP).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
-----
CC      EMBL, X12458; CAA30998.1; -.
CC      EMBL, L44140; AAA92651.1; -.
CC      PIR, S01696; S01696.
CC      MIM, 312090; -.
CC      InterPro: IPR004710; Baa5.
CC      InterPro: IPR002657; BileAc/Na_smptr.
CC      Pfam, PF01758; SBP, 1.
CC      TIGRFAMs, TIGR00841; Baa5; 1.

```

```

SQ      Transmembrane; Transport; Symport.
KW      SEQUENCE   477 AA; 50332 MW; 49CB363EB3B66AID CRC64;

Query Match          16.9%; Score 333.5; DB 1; Length 477;
Best Local Similarity 31.5%; Pred. No. 5.8e-15;
Matches              87; Conservative 53; Mismatches 115; Indels 21; Gaps 3;

QY      12 PANSSEELPGLGEVHGNLELIVFTVSTVMWGLMFSLGSGVEIRKLSHIRRPWGIAVG 71
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       172 PAEDTPATLSADLAHFSENPLIYLPLPIFNKCSF--GCCKVELAEVLKGLMQSPQMLLG 229
QY      72 LLCOFGMLPFPAVYLLAISFSKLPOVALAVLMGCCPGCGTISNITFWVDGMDSIGMTT 131
        :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB       230 LLAGFLVWPPLPAFLIMAKVFMLPKALAGLIITCCSGGGGSYLPSSLGSDVTIAISMTF 289
QY      132 CSTVAALGMPPLCYLTWTWSMSLOONTIPYONIGITLVCTITPVAFGVVNYRPMRKOSK 191
        |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       290 LSTVAATGFLPLSSNAITSRLSIHETLHVPSKLTGLPLFAITANGVLIKSLPRFSQ 349
QY      192 ILIKIGAVGVGVLIL-----VVAAGVVLAKSWMNSDITLTLTSIFPLIGHVTG 241
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       350 LLLGVVPFPFVLLILGGLFLAYRMGFVPLAGIRL-----PIVAGTVPLVGLHW 400
QY      242 FLALLPTHQSQRCTISLETGAONIOMICTMLOLS 277
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       401 YCLATCLKLPVQRRTVISIEGVONSLLALAMLOLS 436

RESULT 9
P3_MOUSE STANDARD; PRI; 182 AA.
AC      P21129;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P3 protein (Fragment).
GN      P3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STAIN-BALB/C; TISSUE=Liver;
RX      MEDLINE=90307023; Pubmed=1973144;
RA      Filippi M., Tribioli C., Tonolo D.;
RT      "Linkage and sequence conservation of the X-linked genes DXS235E (P3)
RT      and DXS254E (Gdx) in mouse and man.";
RL      Genomics 7:453-457(1990).
SC      -1- FUNCTION: THE UBQUITOUS EXPRESSION AND THE CONSERVATION OF THE
CC      SEQUENCE IN DIURNAL ANIMAL SPECIES SUGGEST THAT THE GENE CODES FOR
CC      A PROTEIN WITH HOUSEKEEPING FUNCTIONS.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC      -1- SIMILARITY: BELONGS TO THE SODIUM:BILE ACID SYMPORTER FAMILY
CC      (SBP).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi-sib.ch/announce/
CC      or send an email to license@ebi-sib.ch).
CC      -----
DR      EMBL, J04761; AAA40519.1; -.
DR      InterPro, IPR002657; Billeac/Na_smpctr.
DR      Pfam, PF01758; SBP; 1.
KM      Transmembrane; Transport; Symport.
FT      NON TER
SQ      SEQUENCE   182 AA; 19629 MW; 472D732820CD620 CRC64;

Query Match          9.1%; Score 180.5; DB 1; Length 182;
Best Local Similarity 30.2%; Pred. No. 2.1e-05;
```


Matches 45; Conservative 31; Mismatches 54; Indels 19; Gaps 2;

OY 139 GMPPLCYLWMSLQONLTPYONIGTLVCLTPAFAGVYVYRMPKSKILKIGA 198
 DB 2 GFLPSSAIVSYLSIHETLHPISKILGTLLFPIAAGVIVSKIPKSELLLOVYK 61

OY 199 VVGWGLL-----VVAWAVLAKGSWNSDITLTLTIFPILIGHVGFLLAFT 248
 DB 62 PFSFLLGLFLAHMGVFLVGR-----PIVVGFTVPLVGLVGSIALCL 112

OY 249 HQSWQRCRTISLETGAONIQMCITMLOLS 277
 DB 113 KLPVQQRRTVSIEGVQNSLALAMLOLS 141

RESULT 10
 YCXA_BACSU STANDARD; PRT; 409 AA.
 Q08751;
 01-FEB-1995 (Rel. 31, Created)
 01-FEB-1995 (Rel. 31, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ycxA (ORF5).
 GN YCXA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=93360813; PubMed=8355609;
 RA Cosmina P., Rodriguez F., de Ferra F., Grandi G., Perago M.,
 Venema G., van Sinderen D.;
 RT "Sequence and analysis of the genetic locus responsible for surfactin
 synthesis in *Bacillus subtilis*.";
 RT Mol. Microbiol. 8:821-831(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124189; PubMed=8969502;
 RA Yamane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the *Bacillus subtilis* chromosome:
 determination of the sequence of a 146 kb segment and identification
 of 113 genes.";
 RT Microbiology 142:3047-3056(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriss R., Bourcier L., Brans A., Braun M., Brignell S.C., Broh S.,
 Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 Fritze C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Hentut A.,
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,
 Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,
 Presecan E., Pujic P., Purnelle B., Rappoport G., Rey M., Reynolds S.,
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scofield S.,
 Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambot R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipet A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus
 subtilis*";
 RT Nature 390:249-256(1997).
 RL - SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 CC EMBL; X70356; CAA49820.1; -
 CC EMBL; D50453; BAA08987.1; -
 CC EMBL; Z99105; CAB12147.1; -
 CC EMBL; Z99106; CAB12161.1; -
 CC Subtilisin; BG10172; YCXA.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugax tr; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 9 29
 FT TRANSMEM 49 69 * POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 100 120 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 217 237 POTENTIAL.
 FT TRANSMEM 253 273 POTENTIAL.
 FT TRANSMEM 284 304 POTENTIAL.
 FT TRANSMEM 309 329 POTENTIAL.
 FT TRANSMEM 341 361 POTENTIAL.
 FT TRANSMEM 374 394 POTENTIAL.
 SQ SEQUENCE 409 AA; 44858 MW; 8958A438E7E29D3 CRC64;

Query Match 6.3%; Score 125; DB 1; Length 409;
 Best Local Similarity 20.9%; Pred. No. 0.17;
 Matches 78; Conservative 61; Mismatches 122; Indels 112; Gaps 19;

OY 20 LPVGLVHGNLELVTVSTVM--GLMFSLGCVS---TRKLSHTRRMGAVGLLC 74
 DB 34 LPMADAFHADRSLISVSISIFMTTGVQFVGFIDRFVRKI-----MALGAVC 84

OY 75 ---QGLMPFAVYLLAISFSLKPVQAIYVLMG---CCPGATISNIFFWVDGMDLSI 127
 DB 85 ISASFLVLPYPNVHVS-----AIVGVLGIGYSCAVGVTTQYFISCFDTHKGLAL 137

OY 128 SMTTGSYVAAAGMPPLCYL--YWSWSLQONLTPYONIGTLVCLTP-VAFG--- 179
 DB 138 AILTANAGVIVSPPIMAAPYAGW--OSTVYI---LGIWAAVLPVLVFGMKHP 191

OY 180 -----YVAVYRW-----PKOSKI--ILKIGAVVGVLLLV-----AV 210
 DB 192 PHAQAEVYKKSIDMRGFWNMKQSKLHILVGFVTCFTMGITDAHLVPLIKANVSHV 251

OY 211 AGVVALAKS-----ANSDI-----TLTISFIPPLIGHVGFLLAFTHOS--W- 252
 DB 222 NGMAAFAFAFIIGSLAGWISDLGSRVWLSLIFRLSLICLLPILGHHSDLMY 311

OY 253 -----ORRTISLETGAONIQMCITMLOLSFPAHNVQMLSPPLAYG 294
 DB 312 GFFILLFGLSYTGVPLTPASISSEYQYG-----LIGSLGINFPIHOVAGLSVYAGL 366

OY 295 LFOGLDGLIYVA 307
 DB 367 FFDWTHGYLLIVA 379

RESULT 11

RESULT 13

OPRD_HUMAN STANDARD; PRT; 372 AA.

AC P41143; 01-FEB-1995 (Rel. 31, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Delta-type opioid receptor (DOR-1).

GN OPRD1 OR OPRD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OC NCBI_TaxId=9606;

OX NCBI_TaxId=9606;

RN RN

RP SEQUENCE FROM N.A.

TX TISSUE=Brain cortex, and Striatum;

RA MEDLINE=94260835; PubMed=8201839;

RA Knapp R.J., Malatynska E., Fang L., Li X., Babin E., Nguyen M., Santoro G., Varga E.V., Hruby V.J., Roeseke W.R., Yamamura H.I.; "Identification of a human delta opioid receptor: cloning and expression.";

RT Life Sci. 54:PL463-PL469(1994).

RL [2]

RN RN

RP SEQUENCE FROM N.A.

TX MEDLINE=95107267; PubMed=7808419;

RA Simoni F., Befort K., Gavériaux-Ruff C., Matthes H., Nappay V., Lannes B., Micheletti G., Kleffer B.; "The human delta-opioid receptor: genomic organization, cDNA cloning, functional expression, and distribution in human brain.";

RT Mol. Pharmacol. 46:1015-1021(1994).

RL [3]

RN RN

RP SEQUENCE FROM N.A.

TX Graham D.;

RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

RL [4]

RN RN

RP VARIANT PHE-27.

TX MEDLINE=20434820; PubMed=10982041;

RA Gelernter J., Kranzler H.R.; "Variant detection at the delta opioid receptor (OPRD1) locus and population genetics of a novel variant affecting protein sequence.";

RT Hum. Genet. 107:86-88(2000).

RL CC

CC -!- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. HIGHLY STEREOSELECTIVE. RECEPTOR FOR ENKEPHALINS.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: U07882; AAA18789.1; -

DR EMBL: U10504; AAA83426.1; -

DR EMBL: AL009181; CAAL5671.1; -

DR Genew; HGNC:8153; OPRD1.

DR MIM: 165195; -

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCR_Rhodopsn.

DR PROSITE; PS00237; G PROTEIN RECP_F1_1; 1.

DR PROSITE; PS0262; G PROTEIN RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate; Polymorphism.

KV DOMAIN 1 45 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 46 75 1 (POTENTIAL).

FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 85 102 2 (POTENTIAL).

FT DOMAIN 103 124 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 125 144 3 (POTENTIAL).

FT DOMAIN 145 174 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 175 190 4 (POTENTIAL).

FT DOMAIN 191 215 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 216 238 5 (POTENTIAL).

FT DOMAIN 239 261 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 262 284 6 (POTENTIAL).

FT DOMAIN 285 293 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 294 310 7 (POTENTIAL).

FT DOMAIN 311 372 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DISULFID 121 198 BY SIMILARITY.

FT LIPID 333 333 PALMITATE (POTENTIAL).

FT VARIANT 27 C->F.

FT FT /FTID=VAR_012083.

FT CONFLICT 40 41 AR->PG (IN REF. 1).

FT CONFLICT 348 348 A->P (IN REF. 1).

FT CONFLICT 370 370 A->R (IN REF. 1).

SQ SEQUENCE 372 AA; 40368 MW; 9D483FC93BA2BAE4 CRC64;

Query Match 5.8%; Score 114; DB 1; Length 372;

Best Local Similarity 21.3%; Pred. No. 0.81;

Matches 89; Conservative 61; Mismatches 150; Indels 118; Gaps 20;

QY 3 ANCSS--SACPNSSSEELPVGLEHGNLELVTVSVTWGMLPGLSGVEIRKLS 60

DB 17 ANASDAYSACPSAGANASGPPGARSASLSALAIATA--LYSACAV-- 62

QY 61 HIRPWGIAVGLTQFGIMPFTAVILATIS--FSUKPQVAILAVLIMGCCPGGTSINPTF 117

DB 63 -----GLGNVLVWFGLVTRTKTKTAINTYIFNLADALATSTL--PQSAKYLME 113

QY 118 WVDGMDLSISMTSCSTVAALGMPLCIYLTWSMSLQOULTIPYONIGITLVCLT-- 173

DB 114 WPFGL--LCKAV--LSIDVYNNFTSIFLTMMSV 144

QY 174 ---IPVAGV-YVNTWPKSKIT--LKIGAVGVLLVAVNA---GVVLA-----K 217

DB 145 DRYIAVCHPVALDFRFPAPAKKLINICWVLAGVGAVIMAAVTRPDGAVVCMLOPPS 204

QY 218 GSMNSD---ITLLTISFIPFLIGHVGTFLALFTHOSMORCRITSETGAONIOWC--- 270

DB 205 PSWMDVTKICVPLFAVVPILITVYGLML-----RLKSVRLSSKEKDRSLRR 258

QY 271 ITMLQSTFAHLVQMSFPLAYGLFOLID-----GFLVAAVQYKRLKNKHGKNSGC 326

DB 259 ITRMVLVVVGAFFVCMARIHIFVIWTLVIDRRDPLVVAL-----HLCIALGYANSSL 313

QY 327 TEVCHTRKSTSSRETNALVNEBGAITP-----GPGPMDCRALEPVG--HTSC 376

DB 314 NPVLV-----AFLDENFRKCFRQLCRKPGRGPSSFSRARBATARERTVAC 360

RESULT 14

CYB_TOXGO STANDARD; PRT; 368 AA.

AC 020672; 020928; 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Cytochrome B.

GN COB OR CYTB OR CYB.

OS Toxoplasma gondii.

OG Mitochondrion.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Bimeriida; Sarcocystidae; Toxoplasma.

OC NCBI_TaxId=5811;

OX NCBI_TaxId=5811;

RN RN

RP SEQUENCE FROM N.A.

RA Tournel C., Tomavo S.; "Cytochrome B of Toxoplasma gondii.";

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE OF 10-368 FROM N.A.
RC STRAIN=RH;
RA McFadden D.C., Boothroyd J.C.;
RT Cytochrome B gene from *Toxoplasma gondii*;
Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBQUINOL-CYTOCHROME C REDUCTASE
COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
COUPLED TO ATP SYNTHESIS.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF023246; AAC34138.1; -;
DR InterPro; IPR000179; Cyt b b6.
DR Pfam; PF00032; cytochrome b C; 1.
DR Pfam; PF00033; cytochrome b N; 1.
DR PROSITE; PS00192; CYTOCHROME B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME B_HEME; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
Heme.
FT METAL 82 82 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 96 96 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 178 178 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 192 192 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 368 AA; 41594 MW; CCG76BD3784287CA CRC64;
Query Match 5.7%; Score 112.5; DB 1; Length 368;
Best Local Similarity 24.6%; Pred. No. 1;
Matches 57; Conservative 37; Mismatches 101; Indels 37; Gaps 8;
QY 84 YLLAISFLKPVQALVILM-----GCCPGGTISNIFTFWVDGMDISMTTCTVTAALG 139
DB 34 FLVAMFEVLQITGITLAFRYTSEASCAFAFVVOHLEVAAGMEFRLHATTASFFVFLCI 93
QY 140 MMPLCYLVTWMSLQONLTIPIYONIGITLVCTIVARGVYNNYMPQSKILIKIGAV 199
DB 94 LHMTRGLVWMSYSY---LTTAMMS-GLVLYLLTLTATFAGV-LPMGQMS---FWGAT 144
QY 200 VGGVLLVAVAGVLAAGKSWNSDITL---LTISFTPLIGHVTGFLALFTHQSORCR 256
DB 145 VITNLSPILPYLVPMWLGYYVDVTLKRFVHLFLIPFGICIIIVLHIFVLIHNL----- 199
QY 257 TISLETGAONIQMCTMLQLSF-----TAHVLQMLSPPLAYGFL 298
DB 200 -----GSSNPAGIDTALKAFAFPHMLMTDAKCLSYLIGLIFLQAAGFLMEL 245
RESULT 15
Y944 SYNY3 STANDARD; PRT; 383 AA.
AC P74311;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein slr0944.
OS *Synechocystis* sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
OX NCBI_Taxid=1148;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905221;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Saito S., Kimura T.,
RA Hosouchi T., Matsuo A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
DNA Res. 3:109-116 (1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE ACB3 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D90914; BAA18405.1; -;
DR InterPro; IPR004706; Acb3.
DR Pfam; PF01758; SBF; 1.
DR Pfam; PF01759; SBF; 1.
DR TIGR; TIGR00832; acb3; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 139 159 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 272 292 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
SQ SEQUENCE 383 AA; 42402 MW; 3D8C4CF8EA2FE08B CRC64;
Query Match 5.5%; Score 109.5; DB 1; Length 383;
Best Local Similarity 20.9%; Pred. No. 1.6;
Matches 58; Conservative 55; Mismatches 115; Indels 49; Gaps 13;
QY 40 VMGGLMPSLGSVEIRKLMHIRRPGIAGVLLCQFGIMPTAVLLA---ISFLKPVQ 96
DB 62 ICLFPMWYDINWKIPFQSQARQAVKAPVILTVNNWIKPFTWVIFAQFFGLYFAPLL 121
QY 97 AIAVLIMGCCPGGTISNIF-----TFWVDGMDISMTTCTVTAALGM 140
DB 122 TATEIIRG--QEVTLANSYIAGCILLGAPCTAWVLMW--GLYSYNOGLTVMAVNSL 177
QY 141 MPCLCYLVTWMSL---QONLTIPIYONIGIT-LVCTIPVAFGVYNNY-----RW-PQO 189
DB 178 ANLFLYAPLPGKVLAAANLTVWQIVLSVLIVGLPLAAGYSRWYWLKHKGRQWFSQ 237
QY 190 SKIILKIGAVGVLLLVV--AVAGVLAAGKSWNSDITLTI-----SFIPPLIGHVTG 241
DB 238 FLHYSLPIALVALLTLTLILFAFKGELIVNPLH--IFLIAVPLFIQNFIPL-LITYVLG 294
QY 242 FLIALFTHQSORCRITISLETGAONIQMCTMLQLSF 278
DB 295 LKLKL-----SYDDAAPALIGASNHFEVALATAVMLF 327

Search completed: June 9, 2003, 07:08:36
Job time : 19 secs